Journal of Experimental Child Psychology 111 (2012) 516–533

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Child proportional scaling: Is 1/3 = 2/6 = 3/9 = 4/12?

Ty W. Boyer [a,*], Susan C. Levine [b]

[a] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA
[b] Department of Psychology and Department of Comparative Human Development, University of Chicago, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 27 June 2011
Revised 21 October 2011
Available online 10 December 2011

Keywords:
Proportional reasoning
Relative scaling
Mathematical development
Spatial cognition
Numerical reasoning
Mental transformation

ABSTRACT

The current experiments examined the role of scale factor in children's proportional reasoning. Experiment 1 used a choice task and Experiment 2 used a production task to examine the abilities of kindergartners through fourth-graders to match equivalent, visually depicted proportional relations. The findings of both experiments show that accuracy decreased as the scaling magnitude between the equivalent proportions increased. In addition, children's errors showed that the cost of scaling proportional relations is symmetrical for problems that involve scaling up and scaling down. These findings indicate that scaling has a cognitive cost that results in decreasing performance with increasing scaling magnitude. These scale factor effects are consistent with children's use of intuitive strategies to solve proportional reasoning problems that may be important in scaffolding more formal mathematical understanding of proportional relations.

Introduction

Is splitting four cookies between two people the same as splitting six cookies among three people or eight cookies among four people? Does a mug of coffee with 8 parts coffee and 1 part cream taste the same as a thermos with 32 parts coffee and 4 parts cream? Is a 16-ounce can of soup that costs $2.49 a better bargain than, or the same as, a 24-ounce can of soup that costs $3.74 or a 32-ounce can of soup that costs $4.98? Proportional reasoning requires some understanding of scale relations and arises in everyday problems such as those above as well as in mathematics, science, and engineering (Carpenter, Fennema, & Romberg, 1993; Cramer & Post, 1993; Mazzocco & Devlin, 2008; Mix, Levine, & Huttenlocher, 1999; Pitkethly & Hunting, 1996). As pointed out by Lesh, Post, and Behr

* Corresponding author.
E-mail address: tywboyer@indiana.edu (T.W. Boyer).

0022-0965/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2011.11.001

(1988), proportional reasoning is central to many important topics in the mathematics curriculum that are troublesome for students, including fraction equivalence, long division, place value, percentage calculation, measurement conversions, and derivation of rates. Furthermore, the ability to scale across a wide range of sizes is essential in many areas of science and engineering. To name just a few examples, proportional reasoning is involved in the geosciences to understand the relation between maps and geological features in the real world, in chemistry to solve stoichiometry problems, and in engineering to create and understand scale models used to assess the function and appearance of design concepts. Moreover, the ability to scale relations is involved in problems that involve entities ranging in size from the nano to the cosmic. Yet, we know very little about the ability to scale, even in the visible range, or about how this ability changes over the course of development.

The current study examined the role of scale factor (the multiplicative relation between equivalent proportions) in proportional reasoning problems that involve visually depicted proportional relations. Take, for instance, the cookie splitting example above; four cookies for two people is proportionally equivalent to six cookies for three people, both of which are proportionally equivalent to eight cookies for four people (i.e., all reduce to a 2:1 ratio). The scale factor one derives when comparing these proportions varies, however, as scaling up from a 4:2 ratio to a 6:3 ratio requires a lesser scaling magnitude (i.e., a scale factor of 1.5) than scaling up from a 4:2 ratio to an 8:4 ratio (i.e., a scale factor of 2.0). Furthermore, the directionality of scaling can vary despite equivalent scaling magnitudes. For example, scaling up from a 4:2 ratio to an 8:4 ratio (again, a scale factor of 2.0) is reciprocal to scaling down from an 8:4 ratio to a 4:2 ratio (i.e., a scale factor of 0.50). We asked whether these scaling differences affect children's judgments of proportional equivalence for visually depicted proportional relations.

Piaget and colleagues (Inhelder & Piaget, 1958; Piaget & Inhelder, 1951/1975) argued that proportional reasoning requires an understanding of formal operations, that it emerges late in development, and that success on proportional reasoning tasks during early childhood does not reflect understanding of proportional relations but rather is the product of idiosyncratic strategies and informal naive intuitions. In contrast, more recent approaches characterize the development of mathematical understanding not as "all or none" but rather as moving from partial to more complete, and from more to less contextually dependent, noting the value and effectiveness of early intuitive capacities (Fischbein, 1987; Mix, 2002; Mix, Huttenlocher, & Levine, 2002). According to this view, young children's success on mathematical problems may be based on an inherent number sense (e.g., Dehaene, 1997) and may be of value in scaffolding later, more formal mathematical understandings (e.g., Halberda, Mazzocco, & Feigenson, 2008). Moreover, early intuitive understanding continues to influence our mathematical skills even after the ability to perform formal mathematical operations emerges, as shown by phenomena such as the distance effect (e.g., it is easier to discriminate 2 from 9 than to discriminate 8 from 9) (Dehaene & Akhavein, 1995; Moyer & Landauer, 1967). If this is generally true of mathematical understanding, then children's formal mathematical understanding of proportional relations may be enhanced by instruction that builds on their early intuitive understanding of proportional relations.

In fact, research over the past several decades shows that young children, and perhaps even infants, are sensitive to proportional relations (e.g., Denison & Xu, 2010; Jeong, Levine, & Huttenlocher, 2007; McCrink & Wynn, 2007; Sophian, 2000; Xu & Garcia, 2008). However, these studies also show that there are limits to children's early sense of proportion. For instance, children perform better on problems involving halving and doubling than on problems involving other proportions (Ball, 1993; Spinillo & Bryant, 1991). Furthermore, young children do better when they provide a proportional judgment using an analog scale (e.g., estimate a character's happiness with a given proportion) than when they are asked to make a judgment by choosing among alternative proportional relations (Acredolo, O'Connor, Banks, & Horobin, 1989). Finally, young children perform better on proportional judgment problems when they involve continuous amounts than when they involve discrete quantities (Boyer, Levine, & Huttenlocher, 2008; Jeong et al., 2007; Singer-Freeman & Goswami, 2001; Spinillo & Bryant, 1999; cf. Wing & Beal, 2004, where young children were better at evenly allocating resources in a sharing task when given discrete quantities than when given continuous quantities). When proportional relations involve countable sets, children tend to err by matching on the basis of equivalence of the numerator quantity rather than correctly matching on the basis of the proportional relation (e.g., responding that 2/8 is proportionally equivalent to 2/4 rather than 1/4).

Here we asked whether children are limited by another aspect of proportional reasoning problems, namely, the scale factor between equivalent proportions. We hypothesized that the magnitude of scaling required to compare two proportions, like spatial transformations such as mental rotation, will influence children's ability to correctly judge the proportional equivalence of visually depicted proportional relations. The cognitive cost of scaling may be more apparent when participants use intuitive perceptual processes to solve proportional equivalence problems than when they use more formal mathematical approaches (i.e., division and multiplication) because the former are more likely than the latter to activate visual–perceptual processes. It is possible, however, that a trace of the cost of scaling may remain even when proportional problems are solved using mathematical symbols and operations, similar to what has been shown for numerical comparisons involving arrays of both dots and numerals (e.g., Cantlon, Safford, & Brannon, 2009; Dehaene & Akhavein, 1995).

The impact of scaling magnitude has been largely ignored in the literature on proportional reasoning, but there are a few relevant studies. Consistent with the suggestion that scale factor influences proportional reasoning, Vasilyeva and Huttenlocher (2004) found that 4- and 5-year-olds had greater difficulty in translating a position on a map to a referent space when the map-to-space scalar relation was larger. This was interpreted as evidence that children were accessing a perceptually based form of reasoning, where they transform the metric relations within one space to a space of a different size. Although consistent with our hypothesis that scaling has a cognitive cost, this study investigated only widely discrepant scaling relations and involved only scaling up from a map to a real space and not the reverse (i.e., scaling down from a real space to a map). In another study, Vasilyeva, Duffy, and Huttenlocher (2007) asked children to reproduce a target line depicted within the context of a frame by matching either the target line's absolute extent or its line-to-frame relation. Performance on trials that involved matching the target's line-to-frame relation was generally consistent with our hypothesis, with greater amounts of scaling leading to poorer performance; however, the scalar relation between the target and response frames was not systematically manipulated so as to formally test the cost of scaling.

The experiments we report here systematically manipulated the amount of scaling needed to recognize or produce the proportional relation shown in a target proportion. Both experiments involved computerized tasks that examined young children's ability to match proportions that differ from each other in scale factor. The first experiment used a two-alternative forced-choice format in which children in kindergarten through fourth grade were asked to choose which of two choice alternative proportions matched a target proportion. In the second experiment, children in the same grades were asked to produce a proportion equivalent to a target by manipulating the numerator amount within a frame that specified the denominator. In both experiments, our main prediction was that performance would vary with scale factor, specifically, that proportional equivalence judgment accuracy would decrease as the scale factor between the target and match increased. We also examined the effect of scaling directionality (i.e., scaling up from a smaller target to a larger proportional match vs. scaling down from a larger target to a smaller match), which has not been studied previously to our knowledge. A relatively wide range of ages was used, so as to allow exploration of the emergence of proportional reasoning abilities and examination of whether the effects of scaling apply across a broad development period.

Several other factors also were manipulated to test additional predictions. Foil type was manipulated, and consistent with previous research findings, we predicted that participants would perform better on denominator match foils than on numerator match foils. Problem format also was manipulated, and consistent with previous research, we predicted that performance would generally be superior for participants given stimuli involving continuous quantities than for those given stimuli involving discrete quantities. Moreover, we predicted that scale factor may play a greater role for problems involving continuous quantities than for those involving discrete quantities because young children are more likely to apply intuitive perceptual strategies to continuous format problems and are more likely to apply erroneous counting strategies to discrete format problems (Boyer et al., 2008; Jeong et al., 2007). Finally, because previous research has revealed gender differences for other spatial transformations, notably mental rotation (for reviews, see Halpern et al., 2007; Hyde, Fennema, & Lamon, 1990; Voyer, Voyer, & Bryden, 1995), we included equivalent numbers of boys and girls in each age group so as to examine whether gender differences in proportional scaling exist, perhaps particularly at younger ages, where this type of reasoning is likely to rely on intuitive perceptual skills.

Experiment 1

The first experiment provides an initial investigation of whether scale factor plays a role in kindergarten through fourth-grade children's success on a proportional reasoning task that involves scaling up or scaling down a juice–water mixture with a two-alternative choice task.

Method

Participants

Participants were 161 students (80 girls and 81 boys) recruited from six public schools in a large metropolitan school district in the Midwest region of the United States. There were 33 participants recruited from kindergarten classrooms (16 girls and 17 boys), 32 from first-grade classrooms (16 girls and 16 boys), 33 from second-grade classrooms (16 girls and 17 boys), 32 from third-grade classrooms (16 girls and 16 boys), and 31 from fourth-grade classrooms (16 girls and 15 boys). Mean ages of the children at each grade level were as follows: M_k = 6 years 1 month (6;1), M_1 = 7;0, M_2 = 8;2, M_3 = 9;1, and M_4 = 10;1. All students had written parental consent to participate. Demographic information was recorded at the school level, and on the basis of the available statistics for each school and the number of children tested at each school, we estimate that the sampled population included approximately 28% White, 39% Hispanic, 8% Black, and 19% Asian students, with 6% classified as multiracial or other. Approximately 69% of the children in the sample came from low socioeconomic status (SES) backgrounds (range = 42–96% of students at each school), which was estimated using the percentage of students at each school who were eligible for the free or reduced-cost lunch program as a metric of SES.

Procedure

Participants were given the proportional equivalence task on a laptop computer (Dell Inspiron 1501 with 15-inch screen and 1280 $\times$ 800 resolution), during regular school hours, in rooms adjacent to their classrooms. The experiment was presented with custom developed software programmed in Visual Basic 6.0. During task instructions, the experimenter introduced participants to a teddy bear character named "Wally Bear," whose photo appeared on the screen. The experimenter explained that the character enjoys drinking all kinds of juice, and likes to mix his juice himself, but that he must be careful to have the correct mix of water and juice for each type of mix. During each trial, a small photo of the character appeared on the upper left side of the screen. A target proportion column, composed of colored juice and light blue water parts, appeared below the photo. Two choice alternatives appeared on the right two-thirds of the screen: one correct proportional match and one foil (see Fig. 1). The experimenter asked, "Which of these two [pointing to the two alternatives] is the right mix for the juice Wally Bear is trying to make? Which of these two would taste like Wally Bear's juice?" Participants registered responses by clicking on an icon that appeared below the preferred choice alternative with the computer mouse. After children made a selection, another target proportion and two choice alternatives appeared. The "juice" color on each successive trial was randomly selected from five colors (red, blue, green, yellow, and purple), with the constraint that successive trials always involved a different juice color. There were 24 self-paced trials presented in a random order determined by the computer program. No performance feedback was given.

Experimental design

Each participant completed four problems at each of six different levels of scale factor. There were equal numbers of problems that involved scaling up from a smaller target to a larger choice alternative (e.g., 2/3 to 6/9) and scaling down from a larger target to a smaller choice alternative (e.g., 6/9 to 2/3). Three levels of scaling up (i.e., scale factor = 1.5, 2.0, and 3.0) and three levels of scaling down (i.e., scale factor = 0.67, 0.50, and 0.33) were used (see Table 1). By design, these items were symmetrical around a centralized value with a scale factor of 1.00 (which would represent a "no scaling required" 1:1 identical target match); therefore, scaling up and scaling down trials required the same magnitude of scaling. For example, the scale factors of 0.33 and 3.00 both involve a 3:1 scaling magnitude,

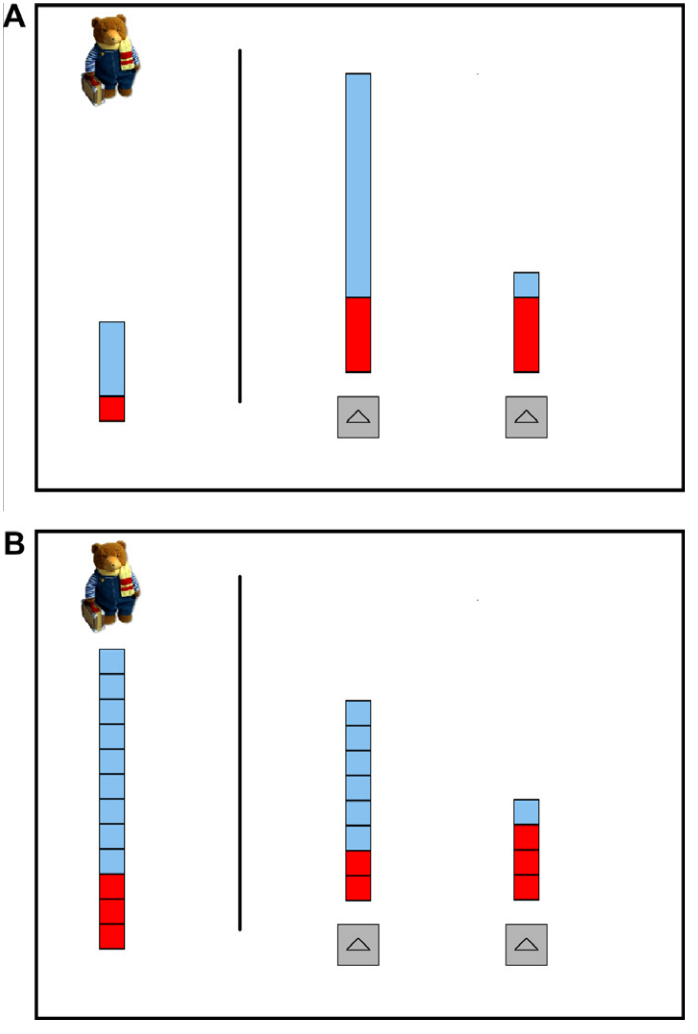

Fig. 1. Study 1: Example screenshots of the proportional reasoning choice task. The target proportion and character photo appear on the left one-third of the screen, and the two choice alternatives appear on the right two-thirds of the screen, each above a corresponding button that participants mouse-clicked to register their response. Panel A illustrates Trial 1 in Table 1, a problem that requires scaling up from the target proportion to the match, and is presented with a denominator foil type and a continuous quantity type. Panel B illustrates Trial 13 in Table 1, a problem that requires scaling down from the target proportion to the match, and is presented with a numerator foil type and a discrete quantity type. Note that in each panel the match is the left choice alternative, but in the experiment the position of the match was randomly selected on each trial.

differing only in scaling direction, with 0.33 involving scaling down from a larger target to a smaller match and 3.00 involving scaling up from a smaller target to a larger match.

All trials involved items that reduced to 1/4, 1/3, 2/3, or 3/4 juice/(juice + water) parts. In addition, the foil alternative was always on the opposite side of the half-boundary from the target proportion. That is, if the target proportion was less than one-half juice/(juice + water), then the foil was greater than one-half and vice versa. This sort of manipulation has been shown to make proportional reasoning tasks easier for younger children (e.g., Spinillo & Bryant, 1991).

Participants were randomly assigned to one of two quantity types, with the constraint of near equal representation of girls and boys in each condition. One group was given problems involving

Table 1
Items used across the 24 experimental trials of Experiment 1.

Trial	Target	Match	Foil	Foil type	Scale factor	Scaling magnitude
Scale factor > 1 (scaling up)						
1	1/4	3/12	3/4	Denominator	3:1 = 3.00	3
2	1/3	3/9	2/3	Denominator	3:1 = 3.00	3
3	2/3	6/9	2/8	Numerator	6:2 = 3.00	3
4	3/4	9/12	3/9	Numerator	9:3 = 3.00	3
5	1/4	2/8	3/4	Denominator	2:1 = 2.00	2
6	1/3	2/6	2/3	Denominator	2:1 = 2.00	2
7	2/3	4/6	2/8	Numerator	4:2 = 2.00	2
8	3/4	6/8	3/9	Numerator	6:3 = 2.00	2
9	2/8	3/12	2/3	Numerator	3:2 = 1.50	1
10	2/6	3/9	2/3	Numerator	3:2 = 1.50	1
11	4/6	6/9	2/6	Denominator	6:4 = 1.50	1
12	6/8	9/12	2/8	Denominator	9:6 = 1.50	1
Scale factor < 1 (scaling down)						
13	3/12	2/8	3/4	Numerator	2:3 = 0.67	1
14	3/9	2/6	3/4	Numerator	2:3 = 0.67	1
15	6/9	4/6	3/9	Denominator	4:6 = 0.67	1
16	9/12	6/8	4/12	Denominator	6:9 = 0.67	1
17	2/8	1/4	2/3	Numerator	1:2 = 0.50	2
18	2/6	1/3	4/6	Denominator	1:2 = 0.50	2
19	4/6	2/3	4/12	Numerator	2:4 = 0.50	2
20	6/8	3/4	2/8	Denominator	3:6 = 0.50	2
21	3/12	1/4	3/4	Numerator	1:3 = 0.33	3
22	3/9	1/3	3/4	Numerator	1:3 = 0.33	3
23	6/9	2/3	3/9	Denominator	2:6 = 0.33	3
24	9/12	3/4	4/12	Denominator	3:9 = 0.33	3

Note. The fractions displayed represent (juice units)/(juice + water units). Scale factor is the match/target ratio, and scaling magnitude captures the direction invariant amount of scaling necessary to go from the target to the match value.

continuous amounts, and the other group was given problems involving discrete units. The only difference was that in the continuous condition the juice and water portions formed unitary columns, whereas in the discrete condition there were lines that demarcated each approximately 1-cm^2 unit (see Fig. 1).

Finally, as can be seen in Table 1 and Fig. 1, there were two different foil types: one that matched the target proportion's colored "juice" part (i.e., the target's numerator; see Table 1) and one that matched the target proportion's "juice + water" whole (i.e., the target's denominator). Previous research suggests that young children have a tendency to match the absolute quantity represented in the target proportion's numerator or denominator rather than match based on the proportional relation (e.g., Boyer et al., 2008; Inhelder & Piaget, 1958). Based on these findings, we purposefully selected foil alternatives that were either numerator or denominator matches. Half of the items involved each foil type, and foil type was nested within scale factor.

Results

The primary analysis was a $3 \times 2 \times 2 \times 5 \times 2 \times 2$ mixed model analysis of variance (ANOVA). Scaling magnitude (1, 2, or 3), scaling direction (scaling up or scaling down), and foil type (numerator match or denominator match) were within-participants factors, and school grade (kindergarten, first, second, third, or fourth), quantity type (continuous or discrete), and child gender (male or female) were between-participants factors. The dependent measure was the number of times participants selected the proportional match. Overall, participants selected the correct proportion on 58.7% of the trials; however, performance varied as a function of scaling magnitude, foil type, and quantity type.

The ANOVA revealed a significant main effect of scaling magnitude, $F(2,282) = 63.37$, $p < .001$, $\eta_p^2 = .310$. As Fig. 2 shows, participants were more likely to select the correct proportional match as the scale factor between the target proportion and the proportional match approached 1.00.

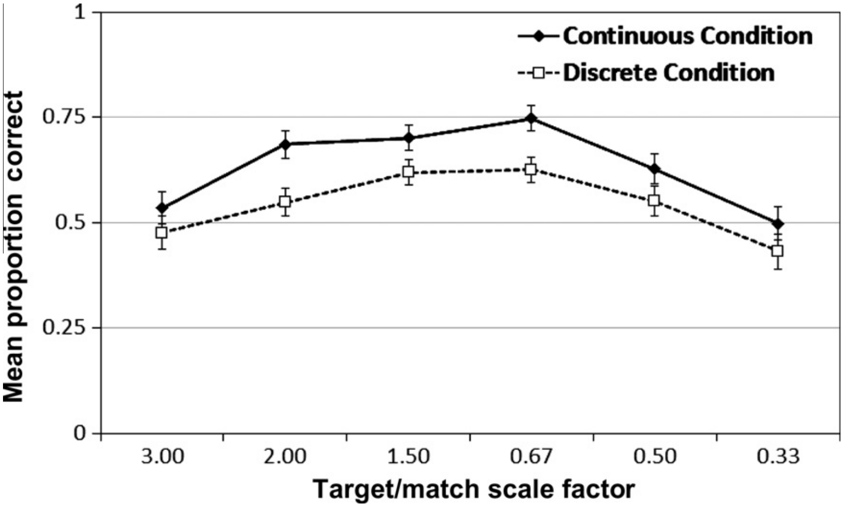

Fig. 2. Study 1: Mean proportion correct at each scale factor for the continuous and discrete conditions, illustrating the main effects of both scaling magnitude and continuity condition and the lack of an effect for scaling direction. Error bars represent standard errors. Note that the *x* axis is scaled to provide equal intervals between each level of scalar relation.

Furthermore, there was no main effect of scaling direction, $F(1,141) < 1$, $p = .45$, $\eta^2_p = .004$. These findings indicate that the difficulty of matching proportions increases as the scaling magnitude increases but do not provide evidence that the difficulty of scaling up differs from that of scaling down. The main effect of foil type was also significant, $F(1,141) = 5.08$, $p = .03$, $\eta^2_p = .04$, with lower accuracy when the foil matched the target's numerator ($M = 56.2\%$) than when the foil matched the target's denominator (i.e., juice + water amount, $M = 61.1\%$). As shown in Fig. 2, participants assigned to the continuous condition ($M = 63.2\%$) performed better than those assigned to the discrete condition ($M = 54.1\%$), $F(1,141) = 5.66$, $p = .02$, $\eta^2_p = .04$. The main effect of school grade also was significant, $F(4,141) = 3.95$, $p = .005$, $\eta^2_p = .10$, reflecting a general increase in performance with school grade ($M_k = 53.7\%$, $M_1 = 50.4\%$, $M_2 = 54.2\%$, $M_3 = 66.4\%$, and $M_4 = 69.4\%$). Third- and fourth-graders chose the proportional match more often than chance, $t(31) = 2.88$, $p = .007$, and $t(30) = 4.18$, $p < .001$, respectively, and did not differ from one another, $t(61) < 1.0$, $p = .69$, whereas the performance of kindergartners, first-graders, and second-graders did not exceed chance (all $ts \leqslant 1.32$, all $ps \geqslant .20$). The main effect of child gender was nonsignificant, $F(1,141) < 1$, $p = .80$, $\eta^2_p < .001$, and none of the interactions involving gender approached statistical significance (all $ps \geqslant .17$).

Several unanticipated interactions were statistically significant. These include the two-way interactions of Foil Type × Scaling Magnitude, $F(2,282) = 9.21$, $p < .001$, $\eta^2_p = .061$, Foil Type × Scaling Direction, $F(1,141) = 5.16$, $p = .025$, $\eta^2_p = .035$, and Scaling Magnitude × Scaling Direction, $F(2,282) = 3.04$, $p = .050$, $\eta^2_p = .021$, and the three-way interaction of Foil Type × Scaling Magnitude × Scaling Direction, $F(8,282) = 4.65$, $p = .010$, $\eta^2_p = .032$. Fig. 3 suggests that these interactions reflect greater symmetry of scaling up and scaling down for denominator match foil problems than for numerator match foil problems. However, post hoc Bonferroni-controlled pairwise comparisons indicate that the only statistically significant difference between scaling up and scaling down occurred between problems that involved doubling and halving (scale factors of 2.0 and .50) for numerator foil problems, $t(160) = 3.47$, $p = .001$, reflecting greater accuracy in scaling up (doubling) than in scaling down (halving) for this pair.

Binomial tests were conducted to examine whether individual participants tended to consistently select the proportional match or the foil alternative. Selecting either choice alternative on at least 18 of the 24 trials was used as a significant difference criterion ($\alpha < .05$, two-tailed binomial test). There were 39 participants (24.2%) who consistently selected the proportional match, compared with 11 participants (6.8%) who consistently selected the foil alternative. In agreement with the group

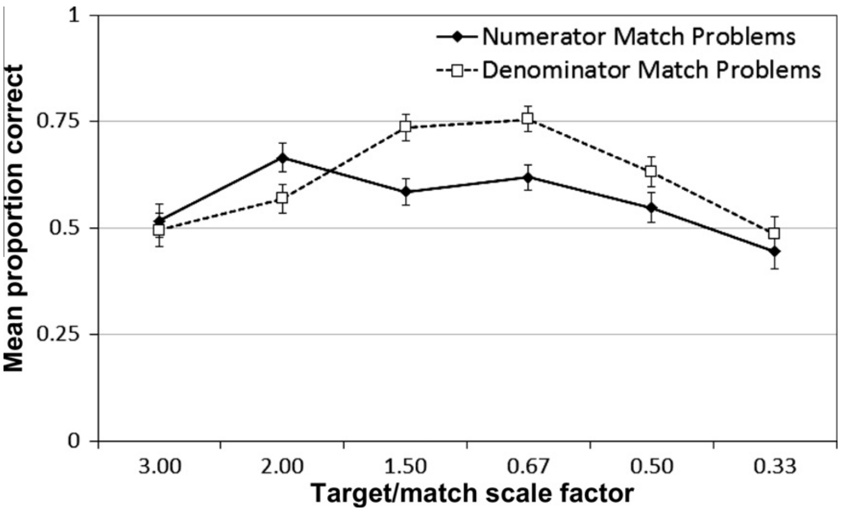

Fig. 3. Study 1: Mean proportion correct at each scale factor for the numerator (i.e., "juice" part) and denominator (i.e., "juice + water" whole) foil types, illustrating the interactions among scaling magnitude, scaling direction, and foil type on the choice task. Error bars represent standard errors.

findings, there was a large jump in the number of participants who consistently selected the proportional match between second and third grade (15.6%, 6.1%, 18.8%, 45.5%, and 35.5% of kindergartners through fourth-graders, respectively). In contrast, 0%, 0%, 15.6%, 15.2%, and 3.2% of kindergarteners through fourth-graders, respectively, consistently selected the foil alternative. There were 23 participants in the continuous condition (28.4%), compared with 16 participants in the discrete condition (20.0%), who consistently selected the proportional match. None of the participants in the continuous condition, but 11 participants in the discrete condition (13.8%), consistently selected the foil alternative, most likely due to using the erroneous strategy of counting the number of units in the target proportion's numerator or denominator and then choosing the foil alternative because it matched this absolute quantity.

Discussion

The current experiment shows that participants were less likely to select the proportionally equivalent choice alternative as scaling magnitude increased. Thus, children's success in perceiving the relational similarity of equivalent proportions decreased as the magnitude of the scale discrepancy between them increased. Consistent with previous studies, participants were more accurate when the problems involved continuous amounts than when they involved discrete sets (Boyer et al., 2008; Duffy, Huttenlocher, & Levine, 2005; Huttenlocher, Duffy, & Levine, 2002; Jeong et al., 2007; Spinillo & Bryant, 1999). However, scale factor affected performance similarly for proportions involving continuous amounts and discrete sets, indicating that the cognitive cost of scaling operates even for problems on which children tend to adopt an erroneous counting strategy.

The effect of school grade was quite clear, with performance levels falling into two grade level groupings; the performance of kindergartners, first-graders, and second-graders was similar, followed by an increase in performance by third- and fourth-graders. This developmental pattern is reflected in terms of overall percentage correct across children as well as in terms of the number of children who consistently respond correctly. In addition, the consistent use of an erroneous counting strategy peaked in second and third grades and dropped off by fourth grade.

The foil alternatives were deliberately manipulated so as to match either the target proportion's numerator (colored juice portion) or its denominator (juice + water whole), thereby providing two kinds of potentially attractive, although erroneous, matches for the target proportion. Consistent with previous studies, accuracy was higher when the foil alternative matched the target's denominator than when it matched the target's numerator (Boyer et al., 2008; Jeong et al., 2007). However, the finding that few kindergartners, first-graders, or second-graders selected either the correct or incorrect choice consistently suggests that they had difficulty in adhering to any specific strategy. Problems involving larger scale relations may have contributed to the inconsistent responding of children, particularly for those in the younger age groups. We suggest that on problems where there was a larger scale factor relation between the target proportion and the proportional match, some children may have thought that the proportional match could not possibly be correct and resorted to selecting the other alternative. In contrast, on problems where there was a smaller scale factor relation between the target and the proportional match, children were more likely to choose the correct proportional match.

One limitation of the current experiment is that the problems involved a simple binary selection. On each problem, participants either were correct or selected the foil alternative provided by the experimenter. Thus, we are unable to assess what was driving children's errors because their response options were restricted to the correct proportional match and an absolute match based on the numerator (half of the problems) or the denominator (half of the problems). Another limitation was that all foil alternatives represent proportional relations on the opposite side of the half-boundary from the correct proportional match, which may enable participants to use "weak" proportional reasoning strategies. That is, participants could choose the correct proportional match by choosing the alternative that has a juice amount (or water amount) on the same side of the half-boundary as the target proportion. Such a strategy would work on the current task but would not generalize to a wide range of proportional reasoning tasks. To address these issues, and to gain a better understanding of the nature of how scale factor contributes to children's difficulties on proportional reasoning problems, Experiment 2 required participants to produce a free response that reflects their estimation of proportional equivalence. Importantly, this method allows a more detailed examination of the misconceptions children may have about proportional equivalence when different scale factors are involved.

Experiment 2

This experiment was designed to further probe the role of scale factor in children's proportional reasoning. The design was similar to that of the previous experiment and made use of the same juice mix cover story. Unlike Experiment 1, however, participants were asked to produce a proportional match in a container that differed in size from the one in which the target proportion was shown. This production task is actually quite similar to the task developed by Vasilyeva and colleagues (2007), where children were asked to replicate the relation of a target line and frame by drawing a line in a different size frame. In our experiment, participants were presented with a target proportion (identical to one of those in the previous experiment) and were asked to adjust the colored juice portion in a second column of a different size until it matched the proportion illustrated by the target.

This procedure has advantages over both the choice procedure used in Experiment 1 and the drawing procedure used by Vasilyeva and colleagues (2007). In particular, the advantage of this task over the alternative choice task used in Experiment 1 is that it provides information about children's construal of what constitutes a proportional match and about the nature of their misconceptions. In addition, it avoids any limitations introduced by the use of foils with proportional relations on the opposite side of the half-boundary from the correct proportional match. Compared with Vasilyeva and colleagues' task, this task provides more structure for participants to produce their responses (i.e., avoiding error that may be due to differences in drawing skill). Moreover, scale factor is explicitly manipulated here to examine whether the scale relation of equivalent proportions influences children's ability to produce proportionally equivalent responses in the context of scaling up and scaling down.

Method

Participants

Participants were 129 students (61 girls and 68 boys) recruited from the same schools as in the previous experiment plus two additional schools. None of the children who participated in Experiment 1 was included in this experiment. There were 27 participants recruited from kindergarten classrooms (13 girls and 14 boys), 24 from first-grade classrooms (12 girls and 12 boys), 24 from second-grade classrooms (12 girls and 12 boys), 27 from third-grade classrooms (10 girls and 17 boys), and 27 from fourth-grade classrooms (14 girls and 13 boys) with mean ages as follows: M_k = 6 years 1 month (6;1), M_1 = 7;0, M_2 = 7;11, M_3 = 8;10, and M_4 = 9;11.

Procedure

As in the first experiment, participants were presented with the task, which was programmed with custom developed software, on a laptop computer. The task instructions were the same as those in the previous experiment. During each trial, Wally Bear's photograph appeared on the upper portion of the left half of the screen, with a target proportion just below it. The response column appeared on the right half of the screen, next to a vertical scroll bar. An orange button with a checkmark on it was just

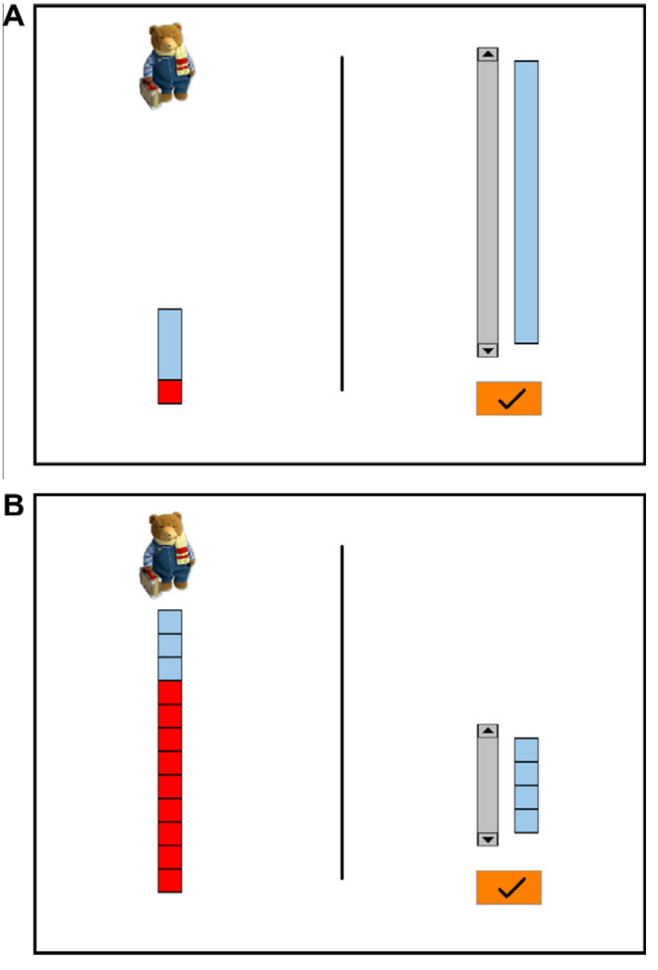

Fig. 4. Study 2: Example screenshots of the proportional reasoning production task. The target proportion and character photo appear on the left half of the screen. The response column appears on the right half of the screen, next to a vertical scroll bar that participants used to adjust the juice portion of the column and above an orange button with a checkmark on it that participants used to register the response. Panel A illustrates Trial 1 in Table 1, a problem that requires scaling up from the target proportion, and is presented with a continuous quantity type. Panel B illustrates Trial 24 in Table 1, a problem that requires scaling down from the target proportion, and is presented with a discrete quantity type.

below the response column and the scroll bar. Fig. 4 provides an example screenshot of what participants were shown during the task.

At the outset of the first trial, the experimenter told participants, "Here [pointing to the target proportion] is Wally Bear's mix for the perfect kind of red juice [note that the actual juice color was randomly selected by the program]. He needs just the right amount of red [pointing to the colored portion] and just the right amount of water [pointing to the light blue "water" portion] so it tastes just right [circling the entire target proportion column]. Now, he wants to make more juice, but he is a little confused and he doesn't know how much red he should use in his juice if he wants to make this much juice [pointing at the response column]. Can you tell him how much red he needs? Look, you can add red by pressing this button right here [pointing to the up arrow at the top of the scroll bar]. You can take some red away by pressing this button right here [pointing to the down arrow at the bottom of the scroll bar]." Participants were then cued to toggle the colored juice portion up and down.

At the start of each trial, the response column was completely composed of light blue "water" units (corresponding to the denominator values in Table 1). As children continuously added or subtracted the juice portion by clicking the scroll bar arrow, the corresponding water amount decreased or increased, respectively. Each click of the scroll bar arrow produced addition or deletion of the colored "juice" portion within the response column of approximately 1.1 mm (such that each 1-cm^2 unit in the discrete condition and the equivalent space in the continuous condition took nine mouse clicks to fill, and responses in both conditions were free to vary continuously between whole units). Children were instructed to press the checkmark button that appeared below the response column when satisfied with the produced proportion. Immediately after this click, another target proportion and another empty production column appeared. The "juice" color changed randomly on each successive trial (as in Experiment 1, five juice colors were used: red, blue, green, yellow, and purple). There were 12 self-paced trials presented in random order.

Experimental design

The design was identical to that of Experiment 1 except that participants were given only half of the problems administered in Experiment 1 (12 rather than 24) due to the greater amount of time required by the production task format. Roughly half of the participants were randomly assigned to receive the target and match column pairs that are listed as the even-numbered trials in Table 1, and the other half were given the odd-numbered pairs. As in Experiment 1, participants were randomly assigned to either the discrete or continuous condition, which again differed in terms of the presence or absence of demarcating lines on the target and response columns.

Results

Error scores were calculated as the difference between the produced juice/whole column proportion and the correct proportional match. For example, a child presented with a 1/4 target proportion (1 juice unit and 3 water units in a 4-unit container) would be asked to reproduce this proportion in a container that is 12 units high (as in Trial 1 in Table 1 and as illustrated in Fig. 4A). If the child produced a response of 5 juice units and 7 water units, this would deviate from the correct proportion by .167 (i.e., 5/12 − 3/12 = 2/12 or .417 − .250 = .167. Note that in calculating errors, responses were rounded to the nearest 1 mm of the produced response, and a whole unit response of 5/12 is used here only for illustrative purposes).[1] The proportional deviation was calculated as the produced proportion minus the target proportion; therefore, when a participant produced a response that was less than the target proportion (i.e., an underestimate), the proportional deviation was negative, and when the participant produced a response that was greater than the target proportion (i.e., an overestimate), the proportional deviation was positive. Although an analysis that includes the sign of the deviation yields findings about whether children tend to systematically undershoot or overshoot the target proportion in their

[1] We also ran the same set of analyses with the absolute physical distance (in millimeters [mm]) from the produced response to the ideal proportional response as the dependent measure. Each of the effects reported here was reflected in this alternative analysis.

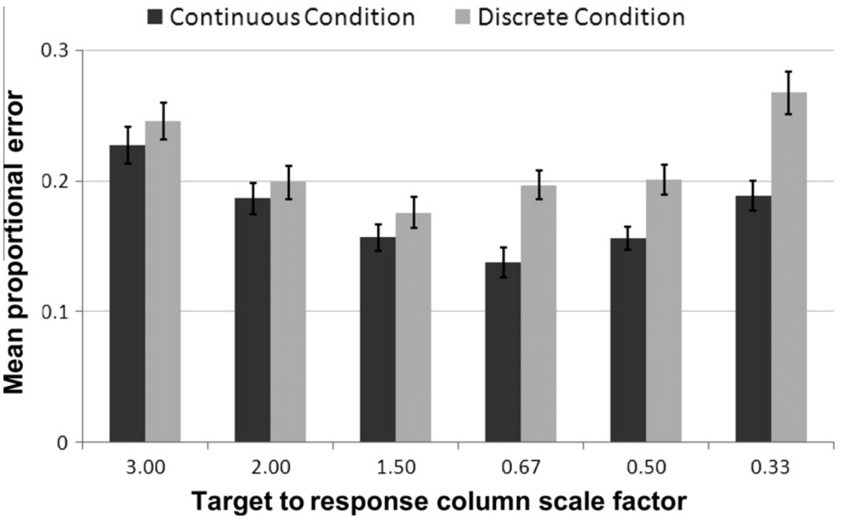

Fig. 5. Study 2: Absolute value of the mean proportional error at each scale factor for the continuous and discrete conditions, illustrating the main effects of scaling magnitude and continuity condition. Error bars represent standard errors.

responses with different scale factors, this kind of coding can also result in a cancelation of errors if they undershoot on some trials and overshoot on others. Therefore, we analyzed children's responses in terms of signed errors in one set of analyses to examine whether there were tendencies to underestimate or overestimate the target proportion and in terms of absolute value of errors in another set of analyses to examine the magnitude of deviations from the correct proportional match. An initial analysis revealed no differences between participants given the odd- or even-numbered item set from Table 1; thus, we did not include this factor in subsequent analyses.

Absolute value error analysis

A mixed model ANOVA was conducted with the mean absolute value of proportional errors at each scale factor as the dependent variable. Scaling magnitude (1, 2, or 3) and scaling direction (scaling up or scaling down) were within-participants factors, and school grade (kindergarten, first, second, third, or fourth), quantity type (continuous or discrete), and child gender (male or female) were between-participants factors. Similar to Experiment 1, Fig. 5 shows that children's accuracy decreased (i.e., mean proportional error increased) as scaling magnitude increased, $F(2,218) = 49.29$, $p < .001$, $\eta_p^2 = .311$. Also similar to Experiment 1, the main effect of scaling direction was not significant, $F(1,109) = 1.66$, $p = .20$, $\eta_p^2 = .015$, suggesting that the effect of scaling magnitude is symmetrical, applying similarly to problems that involved scaling up from a smaller target and to those that involved scaling down from a larger target.

We also found a main effect of quantity type, $F(1,109) = 15.43$, $p < .001$, $\eta_p^2 = .124$, reflecting greater proportional errors in the discrete condition than in the continuous condition (see Fig. 5). Unexpectedly, we found a significant Quantity Type × Scaling Direction interaction, $F(1,109) = 8.77$, $p = .004$, $\eta_p^2 = .074$. Simple main effects analyses reveal that this interaction reflects that the continuous format advantage was significant for problems that involved scaling down from a larger target to a smaller response column, $F(1,109) = 25.20$, $p < .001$, $\eta_p^2 = .188$, but did not reach significance for problems that involved scaling up from a smaller target to a larger response column, $F(1,109) = 2.51$, $p = .12$, $\eta_p^2 = .023$; however, as in scaling down problems, there was a trend toward an advantage for the continuous condition. The analysis also revealed a significant effect of school grade, $F(4,109) = 6.17$, $p < .001$, $\eta_p^2 = .184$, similar to the school grade effect found in Experiment 1. Kindergartners showed the greatest proportional errors ($M = .225$), with a decrease at each subsequent school grade ($M_1 = .212$, $M_2 = .206$, $M_3 = .186$, and $M_4 = .153$). Neither the main effect of child gender, $F(1,109) < 1$, $p = .88$, $\eta_p^2 < .001$, nor any interaction involving child gender was significant (all $ps \geqslant .12$).

Signed error analysis

We next carried out a repeated measures ANOVA with the mean signed proportional error at each scale factor (3.0, 2.0, 1.5, 0.67, 0.50, and 0.33) as the dependent measure. This analysis examined whether responses systematically overestimated or underestimated the correct proportion at different scale factors. There was a main effect of scaling magnitude, $F(5,640) = 8.34$, $p < .001$, $\eta_p^2 = .06$, such that signed proportional errors increased with scaling magnitude, as was the case for absolute value of proportional errors. Furthermore, participants tended to underestimate the target proportion when scaling up (i.e., scale factors 3.0, 2.0, and 1.5) and overestimate the target proportion when scaling down (i.e., scale factors .67, .50, and .33). This effect was best fit by a linear trend, $F(1,128) = 12.49$, $p = .001$, $\eta_p^2 = .09$.

Demarcating line responses

A final set of analyses was conducted to more closely examine whether participants in the discrete quantity type condition tended to produce responses that matched one of the lines that demarcated the response column rather than produce responses that fell in between these demarcating lines. As a baseline comparison, we compared this potential "demarcating line response bias" with the tendency to produce responses that corresponded to demarcating line locations in the continuous quantity type condition. A majority of the responses of children in the discrete condition ($M = 78.9\%$ of responses) matched one of the demarcating lines on the response column (i.e., the response was within 1 mm of one of the lines). Of these responses, only 16.4% corresponded with the correct proportional match line, 24.4% were at the line specifying the absolute target numerator quantity, 10.8% were at the absolute target denominator quantity, and 48.5% were at one of the other lines. In contrast, in the continuous quantity type condition, a much smaller proportion of responses occurred at locations corresponding to the demarcating lines ($M = 30.5\%$). Moreover, 23.7% of these corresponded with the location of the correct proportional match, 26.7% corresponded with the absolute target numerator quantity, 6.7% corresponded with the absolute target denominator quantity, and 42.8% were where one of the other lines would have been.

An ANOVA, with quantity type and school grade analyzed between participants, was conducted to examine differences in the rates of producing responses that matched one of the demarcating lines. This revealed a significant main effect for quantity type, $F(1,119) = 199.24$, $p < .0001$, $\eta_p^2 = .63$, indicating that responding on a demarcating line in the discrete condition was significantly more frequent than responding at a location corresponding to these lines in the continuous condition. The main effect for school grade was nonsignificant, $F(4,119) = 0.34$, $p = .85$, $\eta_p^2 = .01$, as was the School Grade $\times$ Quantity Type interaction, $F(4,119) = 0.69$, $p = .60$, $\eta_p^2 = .02$.

One last set of tests revealed that in the discrete condition the mean absolute value of proportional errors was higher for responses that fell on one of the demarcating lines (including the minority of responses that corresponded with the line designating the correct proportion) than for responses that did not fall on one of the demarcating lines, $t(65) = 2.74$, $p = .008$. There was no such difference for the corresponding locations in the continuous condition, $t(60) = 0.37$, $p = .71$. Finally, comparing quantity type conditions, the mean absolute value of proportional errors for responses that matched one of the lines in the discrete condition was significantly higher than that for responses in the corresponding locations in the continuous condition, $t(127) = 3.66$, $p < .001$. In contrast, there was no significant difference between the continuous and discrete conditions when responses did not fall on one of the demarcating lines, $t(127) = 0.51$, $p = .61$. These results clearly suggest that participants in the discrete condition tended to adhere to the unit structure implied by the demarcating lines, which was a strategy not available to the participants in the continuous condition, and that this erroneous strategy led them astray on the task and was associated with a decrease in their overall performance.

Discussion

The findings of the current experiment converge with the findings of Experiment 1. That is, participants performed better in the continuous condition than in the discrete condition and produced progressively larger errors as the scale relation between the target proportion and the correct response proportion increased. These results emerged both when responses were analyzed in terms of the absolute value of the errors and when they were analyzed in terms of directional errors. Moreover, these results were found for both continuous format proportions and discrete format proportions, and they were generally true for both scaling up and scaling down.

In interpreting the significant scale factor effect, we considered the possibility that this finding could reflect a tendency of children to match the absolute juice amount of the target, a tendency that would result in a spurious scaling effect. For example, in scaling up from a target of 1/4 to a response of 2/8, a participant who consistently tried to match the juice part of the target would respond 1/8, resulting in an error of 2/8 – 1/8 or .125. In contrast, in scaling up from a target of 1/4 to 3/12, a participant who consistently matched the absolute target quantity would respond 1/12, resulting in a larger error of 3/12 – 1/12 or .167. Several aspects of the data, however, suggest that this possible response tendency does not account for the scale factor effect. First, the results of the production task used in Experiment 2 were consistent with the results of the forced-choice task used in Experiment 1, where for "whole" foil-type problems (i.e., foil choice matched the target's denominator) this tendency would not account for the scaling effect. Moreover, in Experiment 1, a Scale Factor × Foil Type interaction actually showed that scale factor had a more pronounced effect on the "whole" foil type problems than on the "part" foil type problems (i.e., foil choice matched the target's numerator). Second, the signed proportional error analysis used in Experiment 2 showed that on average responses fell between the absolute match amount and the ideal proportion. Thus, even if participants were influenced by the absolute target (i.e., "juice") amount, they were (at least on average) adjusting their response in the appropriate direction. This pattern of responding suggests that participants were influenced by the absolute juice amount in the target proportion as well as by the relative proportion of juice/water in the target proportion.

The continuity condition effects were largely consistent with expectations. In general, error rates were higher in the discrete condition than in the continuous condition. Examined more closely, the data indicate that participants in the discrete condition produced responses that were on one of the demarcating lines a sizable majority of the time. This indicates that participants were attending to the units defined by these lines; however, because error rates of participants in the discrete condition were consistently higher than those of participants in the continuous condition, specifically for responses that fell on demarcating lines, the demarcating line bias actually amplified error rates. That is, whereas children in the continuous condition were more likely to be "in the ballpark" of the target proportion, participants in the discrete condition appear to have been, in many instances, led astray by the structure provided by the demarcating lines. On the majority of trials, participants produced responses that fell on the wrong line rather than on the correct line, resulting in increased proportional error relative to the continuous condition. This tendency may be related to children's difficulty in unitizing continuous quantities (Huntley-Fenner, 2001; Sophian, 2007). This is also particularly interesting because adults presumably would exploit the additional precision that the demarcating lines afford and would be even more likely to produce responses with minimal proportional error in the discrete condition than in the continuous condition. However, children's developing understanding of number, coupled with a commitment to the idea that counting units is important in many problem-solving situations, actually contributed to a decrement in proportional reasoning, as other studies have shown previously (e.g., Boyer et al., 2008).

The developmental effects in the current experiment showed that third-graders, and especially fourth-graders, produced the target proportions more accurately than the younger groups (i.e., kindergartners, first-graders, and second-graders). However, the results also suggest that children in each of the grade levels included (kindergarten through fourth grade) were influenced by both the correct proportion and the absolute juice amount, with a shift toward producing responses closer to the proportional match with increasing school grade.

General discussion

Our findings indicate that young children's ability to reason about the equivalence of visually depicted proportions is limited by the scale factor that relates equivalent proportions. In Experiment 1 participants were more likely to select a choice alternative that was proportionally equivalent to a target proportion if less scaling was required, and in Experiment 2 participants were more accurate at producing a response near the target proportion if less scaling was required. These findings extend prior research showing that young children are more successful in relating a map to a real space if the scalar relation between the two is closer to 1:1 (Vasilyeva & Huttenlocher, 2004).

What factors might contribute to the scale factor effect? One possibility is that children solve these problems by mentally expanding or shrinking the target proportion for problems that involve scaling up or scaling down, respectively. Larger transformations may result in greater cognitive costs and decreased response accuracy, analogous to what has been demonstrated for spatial tasks such as mental rotation (Kail, 1985; Shepard & Metzler, 1971) or mentally traversing a distance (Kosslyn, 1994). If children are using this sort of analog mental transformation to relate proportions that differ in terms of scale factor, an advantage for smaller scaling magnitudes would be expected, just as has been found for mental rotation problems involving smaller angular disparities.

We also considered and rejected several other possible explanations for the scale factor effect. One was the tendency to match the number of units in the numerator. Although this could potentially account for the findings in Experiment 2, it could not account for the scale factor effect found in denominator foil trials in Experiment 1, which was actually significantly larger than the scale factor effect found for numerator foil trials in that experiment. Nonetheless, it is possible that this response tendency contributed to the finding of a scale factor effect in the discrete conditions in Experiments 1 and 2.

It is also possible that explicit mathematical problem-solving strategies could lead to scale factor effects if the difficulty of the calculations involved in figuring out proportional equivalence is greater for problems involving larger scale factors than for those involving smaller scale factors. For instance, if problems involving lower scaling magnitudes were composed of a smaller number of items than problems with larger scaling magnitudes, it would be reasonable to suppose that participants, particularly younger participants, would have more difficulty in solving those problems with greater scaling magnitudes due to limitations related to their mathematical knowledge. To assess this possibility, we examined the problems presented, revealing that problems involving smaller scale factors (i.e., those with 1.5 and 0.67 scale factors) were actually more mathematically difficult than those involving larger scale factors (i.e., those with 3.0 and 0.33 scale factors). This holds in terms of the total number of units involved (on average, there were 7.5, 5.5, and 6.67 units in each proportion for problems with scaling magnitudes of 1, 2, and 3, respectively; see Table 1). This also applies in terms of whether the problems involved whole number or fractional multiplicative relations. For instance, whereas the match in Trial 1, which involves a large scale factor, is readily multiplied by the target value (i.e., the match is $3\times$ the target value), the mathematical operation required in Trial 9, which involves a smaller scale factor, is not so simple (i.e., the match is $1.5\times$ the target value). Yet in both experiments, performance on Trial 9 is better than performance on Trial 1 (in Experiment 1 the mean accuracy was 61% on Trial 9 vs. 47% on Trial 1, and in Experiment 2 the mean proportional error was .18 on Trial 9 vs. .22 on Trial 1).

Thus, the scaling magnitude effect reported here seems to be most consistent with the possibility that young children's proportional reasoning is dependent on the ease of carrying out the analog mental transformation required to relate the two visually depicted proportions. This interpretation aligns with previous research showing that children can solve proportional reasoning problems by relying on intuitive strategies (Ahl, Moore, & Dixon, 1992; Boyer et al., 2008; Empson, Junk, Dominguez, & Turner, 2005; Falk & Wilkening, 1998; Mack, 1993; Sophian, 2000). The fact that all age groups included in our experiments showed the scale factor effect, with no significant decrease in the magnitude of this effect with age, raises the question of whether traces of the scale factor effect would remain even for adults, who have the capacity to use more formal mathematics to derive their answers. This persistence of early intuitive mathematics would be consistent with other findings in the literature such as the finding that the ratio dependency of the approximate number system influences our numerical judgments throughout the lifespan even when we are making judgments that involve number symbols (e.g., Dehaene, 1997; Dehaene & Akhavein, 1995). Moreover, findings showing that the acuity of the approximate number system is correlated with later mathematics achievement levels (e.g., Halberda & Feigenson, 2008; Halberda et al., 2008) raise the possibility that accuracy in intuitive proportional judgments may also be related to individual differences in formal mathematical skills.

Prior research also indicates that linear number line representation predicts more general mathematics proficiency (Booth & Siegler, 2006; Laski & Siegler, 2007) as well as performance on spatial and numerical reasoning tasks (Opfer, Thompson, & Furlong, 2010). Similarly, early intuitive proportional scaling may be related to children's mathematics aptitude. However, as Thompson and Opfer (2008) reported, linear representations of number may be maladaptive for the representations of fractions on the number line. It remains an open question as to how early representations of number, proportional relations, and other skills, such as spatial visualization skills, combine to predict later achievement in mathematics.

As has been reported previously, children's proportional reasoning was better in the context of continuous amounts than in the context of discrete units (Jeong et al., 2007; Singer-Freeman & Goswami, 2001; Spinillo & Bryant, 1999). This suggests that children's intuitions about proportion are masked when they are able to adopt the strategy of counting and matching in terms of absolute quantity to arrive at answers to proportional reasoning problems. Furthermore, the production task used in Experiment 2 demonstrates that the subtle differences in quantity type (i.e., the presence vs. absence of demarcating lines) played a significant role in the responses participants produced given that participants in the discrete condition had a marked tendency to produce responses that coincided with the lines that demarcated the discrete units. Furthermore, this demarcating line response bias was associated with less accurate performance, consistent with previous research showing that counting units interferes with proportional reasoning (e.g., Boyer et al., 2008; Jeong et al., 2007; Singer-Freeman & Goswami, 2001; Spinillo & Bryant, 1999).

It is important to point out that although the effect size of scaling magnitude, the variable of particular interest in the current study, was moderate (i.e., $\eta_p^2 = .31$ in both experiments), the effect sizes of several other factors, including the continuous versus discrete quantity effect, were small (i.e., many of the η_p^2 values that we reported were <.10). It is possible that the continuous versus discrete quantity type effect size was reduced by the inclusion of problems that involved large scale factors on which performance was poor independent of quantity type.

Finally, and not surprisingly, in both experiments we found that performance level improved with age. In Experiment 1 there was a relatively dramatic increase in performance between second and third grade (i.e., between ~8 and 9 years of age), and in Experiment 2 the more dramatic increase took place slightly later, between third and fourth grade (i.e., between ~9 and 10 years of age), perhaps because of the greater demands of the production task. We suggest that children in the early elementary school grades are in a transition period characterized by uncertainty about which strategies are best to solve proportional reasoning problems. Prior to third or fourth grade, children have little exposure to the algorithms needed to calculate proportions (Fuson & Abrahamson, 2005; Pitkethly & Hunting, 1996). In the absence of this instruction, children seem to rely either on an intuitive strategy to match the target proportion (more common at small scale factors and with continuous amounts) or on erroneous counting strategies (more common at large scale factors and with discrete units). Whereas younger children have a tendency to err by overrelying on their limited understanding of number and its utility (e.g., counting and matching absolute quantities and relying on the demarcating unit lines to produce responses), older children are more likely to have some understanding of proportional relations and are better able to inhibit tendencies to employ numerical matching strategies.

Our findings indicate that scaling visually depicted proportional relations carries a cognitive cost that increases with the magnitude of the scaling required. In this sense, as discussed above, early proportional reasoning may engage mental transformations that are analogous to other spatial tasks (e.g., mental rotation). In terms of practical implications, the findings suggest that it may be effective to introduce proportional reasoning to young children using examples that involve continuous amounts and small scale factors. The initial use of such problems in lessons on proportion and the gradual extension to problems involving discrete quantities and larger scale factors might allow students to connect the formal mathematical operations they are learning to their earlier emerging intuitions and lead to earlier and deeper understandings of proportionality (Fischbein, 1987). An essential next step is to identify the best ways to use early intuitive notions of proportionality to scaffold more explicit understandings. This is an important future direction given the centrality of proportional reasoning in mathematics and science and the difficulty that many children experience with this topic.

Acknowledgments

This research was supported by a grant from the National Science Foundation (Spatial Intelligence and Learning Center Grant, SBE0541957). The authors thank Kevin Dietz for assistance with data collection, the Chicago Public School System, the school administrators and teachers who facilitated the research, and the students who participated.

References

Acredolo, C., O'Connor, J., Banks, L., & Horobin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology. *Child Development, 60*, 933–945.
Ahl, V. A., Moore, C. F., & Dixon, J. A. (1992). Development of intuitive and numerical proportional reasoning. *Cognitive Development, 7*, 81–108.
Ball, D. (1993). Halves, pieces, and twoths: Constructing and using representational contexts in teaching fractions. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 157–195). Hillsdale, NJ: Lawrence Erlbaum.
Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*, 189–201.
Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology, 44*, 1478–1490.
Cantlon, J. F., Safford, K. E., & Brannon, E. M. (2009). Spontaneous analog number representations in 3-year-old children. *Developmental Science, 13*, 289–297.
Carpenter, T. P., Fennema, E., & Romberg, T. A. (1993). *Rational numbers: An integration of research*. Hillsdale, NJ: Lawrence Erlbaum.
Cramer, K., & Post, T. (1993). Connecting research to teaching proportional reasoning. *Mathematics Teacher, 86*, 404–407.
Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and level of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 314–326.
Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*, 798–803.
Duffy, S., Huttenlocher, J., & Levine, S. (2005). Is it all relative: How young children encode extent. *Journal of Cognition and Development, 6*, 51–63.
Empson, S. B., Junk, D., Dominguez, H., & Turner, E. (2005). Fractions as the coordination of multiplicatively related quantities: A cross-sectional study of children's thinking. *Educational Studies in Mathematics, 63*, 1–28.
Falk, R., & Wilkening, F. (1998). Children's construction of fair chances: Adjusting probabilities. *Developmental Psychology, 34*, 1340–1357.
Fischbein, E. (1987). *Intuition in science and mathematics*. Dordrecht, The Netherlands: Reidel.
Fuson, K. C., & Abrahamson, D. (2005). Understanding ratio and proportion as an example of the apprehending zone and conceptual-phase problem-solving models. In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 213–234). New York: Psychology Press.
Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*, 1457–1465.
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with mathematics achievement. *Nature, 455*, 665–668.
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–51.
Huntley-Fenner, G. (2001). Why count stuff? Young preschoolers do not use number for measurement in continuous dimensions. *Developmental Science, 4*, 456–462.
Huttenlocher, J., Duffy, S., & Levine, S. (2002). Infants and toddlers discriminate amount: Are they measuring? *Psychological Science, 13*, 244–249.
Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139–155.
Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
Jeong, Y., Levine, S., & Huttenlocher, J. (2007). The development of proportional reasoning: Effect of continuous vs. discrete quantities. *Journal of Cognition and Development, 8*, 237–256.
Kail, R. (1985). Development of mental rotation: A speed–accuracy study. *Journal of Experimental Child Psychology, 40*, 181–192.
Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 76*, 1723–1743.
Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 93–118). Reston, VA: National Council of Teachers of Mathematics.
Mack, N. K. (1993). Learning rational numbers with understanding: The case of informal knowledge. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 85–106). Hillsdale, NJ: Lawrence Erlbaum.
Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and "holes": Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science, 11*, 681–691.
McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science, 18*, 740–745.
Mix, K. S. (2002). The construction of number concepts. *Cognitive Development, 17*, 1345–1363.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood*. New York: Oxford University Press.
Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology, 35*, 164–174.
Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*, 1519–1520.
Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial–numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science, 13*, 761–771.
Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: Norton [Original work published 1951].
Pitkethly, A., & Hunting, R. (1996). A review of recent research in the area of initial fraction concepts. *Educational Studies in Mathematics, 30*, 5–38.
Shepard, R., & Metzler, J. (1971). Metal rotation of three-dimensional objects. *Science, 171*, 701–703.
Singer-Freeman, K. E., & Goswami, U. (2001). Does a half pizza equal half a box of chocolate? Proportional matching in an analogy task. *Cognitive Development, 16*, 811–829.
Sophian, C. (2000). Perceptions of proportionality in young children: Matching spatial ratios. *Cognition, 75*, 145–170.
Sophian, C. (2007). *The origins of mathematical knowledge in childhood*. Mahwah, NJ: Lawrence Erlbaum.
Spinillo, A. G., & Bryant, P. (1991). Children's proportional judgments: The importance of "half". *Child Development, 62*, 427–440.
Spinillo, A. G., & Bryant, P. (1999). Proportional reasoning in young children: Part–part comparisons about continuous and discontinuous quantity. *Mathematical Cognition, 5*, 181–197.
Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effect of context on age and sex differences in magnitude estimation. *Journal of Experimental Child Psychology, 101*, 20–51.
Vasilyeva, M., Duffy, S., & Huttenlocher, J. (2007). Developmental changes in the use of absolute and relative information: The case of spatial extent. *Journal of Cognition and Development, 8*, 455–471.
Vasilyeva, M., & Huttenlocher, J. (2004). Early development of scaling ability. *Developmental Psychology, 40*, 682–690.
Voyer, D., Voyer, S., & Bryden, M. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.
Wing, R. E., & Beal, C. R. (2004). Young children's judgments about the relative size of shared portions: The role of material type. *Mathematical Thinking and Learning, 6*, 1–14.
Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 5012–5015.